問6.1

群$G_1 \ni x_i^{(1)}$のy軸上での(不偏)分散は

$$\frac{1}{n_1-1}\sum_{i=1}^{n_1}\left(y_i^{(1)}-\bar{y}^{(1)}\right)^2 = \frac{1}{n_1-1}\sum_{i=1}^{n_1}\left\{w_1^2\left(x_{i1}^{(1)}-\bar{x}_1^{(1)}\right)^2 + 2w_1w_2\left(x_{i1}^{(1)}-\bar{x}_1^{(1)}\right)\cdot\left(x_{i2}^{(1)}-\bar{x}_2^{(1)}\right) + w_2^2\left(x_{i2}^{(1)}-\bar{x}_2^{(1)}\right)^2\right\}$$

$$\boxed{\bar{y}^{(1)} = \frac{1}{n_1}\sum w^T x_i^{(1)}}$$

$$= w_1^2 S_{11}^{(1)} + 2w_1w_2S_{12}^{(1)} + w_2^2S_{22}^{(1)}$$

$$\boxed{\begin{array}{l} S_{11}^{(1)} = \frac{1}{n_1-1}\sum_i\left(x_{i1}^{(1)}-\bar{x}_1^{(1)}\right)^2,\quad S_{12}^{(1)} = S_{21}^{(1)} = \frac{1}{n_1-1}\sum_i\left(x_{i1}^{(1)}-\bar{x}_1^{(1)}\right)\left(x_{i2}^{(1)}-\bar{x}_2^{(1)}\right) \\ S_2^{(1)} = \frac{1}{n_1-1}\sum_i\left(x_{i2}^{(1)}-\bar{x}_2^{(1)}\right)^2 \end{array}}$$

$$= (w_1, w_2)\begin{pmatrix} S_{11}^{(1)} & S_{12}^{(1)} \\ S_{21}^{(1)} & S_{22}^{(1)} \end{pmatrix}\begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$$

$$= w^T S_1 w \quad \blacksquare$$

問6.2

λをwの関数と考えると、$\alpha \in \mathbb{R}$を用いて

$$\lambda(\alpha w) = \frac{\left\{(\alpha w)^T(\bar{x}_1-\bar{x}_2)\right\}^2}{(\alpha w)^T S (\alpha w)} = \frac{\left\{w^T(\bar{x}_1-\bar{x}_2)\right\}^2}{w^T S w} = \lambda(w) \quad \text{—— (☆)}$$

となりwのスケールによらないことが分かる。そこで$w^T S w = 1$とおけば固有値問題:
$w^T S w - 1 = 0$の下で$w^T(\bar{x}_1-\bar{x}_2)$を最大化させる、つまり$\left\{w^T(\bar{x}_1-\bar{x}_2)\right\}^2 - \lambda(w^T S w - 1) = 0$
の極値条件: $\frac{\partial L}{\partial w} = 2(\bar{x}_1-\bar{x}_2)w - 2\lambda S w = 0$をみたす$w$を求める問題に帰着する。
(ラグランジュの未定乗数法)。いま、$(\bar{x}_1-\bar{x}_2)w$とλがスカラーであるから

$w \propto S^{-1}(\bar{x}_1-\bar{x}_2)$となる。$\lambda$は$w$に関して定数倍の自由度がある(☆)から

$\hat{w} = S^{-1}(\bar{x}_1-\bar{x}_2)$を得る。

問6.3

(1) $S^{-1} = \frac{1}{0.24-0.04}\begin{pmatrix} 0.6 & 0.2 \\ 0.2 & 0.4 \end{pmatrix} = \begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}$ より

$$D_1^2 = (\boldsymbol{x}-\bar{\boldsymbol{x}}_1)^\top S^{-1} (\boldsymbol{x}-\bar{\boldsymbol{x}}_1)$$
$$= (x_1, x_2-1)\begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2-1 \end{pmatrix}$$
$$= (x_1, x_2-1)\begin{pmatrix} 3x_1+x_2-1 \\ x_1+2x_2-2 \end{pmatrix}$$
$$= 3x_1^2 + x_1x_2 - x_1 + x_1x_2 - x_1 + 2x_2^2 - 2 - 2x_2 + 2$$
$$= \underline{3x_1^2 + 2x_1x_2 + 2x_2^2 - 2x_1 - 2x_2 + 2}$$

$$D_2^2 = (\boldsymbol{x}-\bar{\boldsymbol{x}}_2)^\top S^{-1} (\boldsymbol{x}-\bar{\boldsymbol{x}}_2)$$
$$= (x_1-1, x_2)\begin{pmatrix} 3 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x_1-1 \\ x_2 \end{pmatrix}$$
$$= (x_1-1, x_2)\begin{pmatrix} 3x_1+x_2-3 \\ x_1+2x_2-1 \end{pmatrix}$$
$$= 3x_1^2 - 3x_1 + x_1x_2 - x_2 - 3x_1 + 3 + x_1x_2 + 2x_2^2 - x_2$$
$$= \underline{3x_1^2 + 2x_1x_2 + 2x_2^2 - 6x_1 - 2x_2 + 3} \ //$$

(2) $h(\boldsymbol{x}) = D_2^2 - D_1^2 = -4x_1 + 1$

(3) $h((0.5, 0.8)^\top) = -1 \leq 0$ より D_2 に属する。

問6.4

(6.34)式を(6.33)式に代入すると、

$$\lambda = \frac{\{(S^{-1}(\bar{\boldsymbol{x}}_1-\bar{\boldsymbol{x}}_2))^\top (\bar{\boldsymbol{x}}_1-\bar{\boldsymbol{x}}_2)\}^2}{(S^{-1}(\bar{\boldsymbol{x}}_1-\bar{\boldsymbol{x}}_2))^\top S\, S^{-1}(\bar{\boldsymbol{x}}_1-\bar{\boldsymbol{x}}_2)}$$

$$= \frac{\{(\bar{\boldsymbol{x}}_1-\bar{\boldsymbol{x}}_2)^\top S^{-1}(\bar{\boldsymbol{x}}_1-\bar{\boldsymbol{x}}_2)\}^2}{(\bar{\boldsymbol{x}}_1-\bar{\boldsymbol{x}}_2)^\top S^{-1}(\bar{\boldsymbol{x}}_1-\bar{\boldsymbol{x}}_2)} \quad (\because S^{-1}\text{は対称行列})$$

$$= (\bar{\boldsymbol{x}}_1-\bar{\boldsymbol{x}}_2)^\top S^{-1}(\bar{\boldsymbol{x}}_1-\bar{\boldsymbol{x}}_2) \ \blacksquare$$

問6.5 群の重心からのキョリがユークリッド距離で等しい異なる2点は、マハラノビス距離では確率楕円の等高線の高さの差がキョリに相当してキョリが異なる。

問6.6

(1) 標本平均: $\bar{y}^{(j)} = \frac{1}{n_j} \sum_{i=1}^{n_j} y_i^{(j)} = \frac{1}{n_j} \sum_{i=1}^{n_j} w^T x_i^{(j)} = w^T \bar{x}_j$

〃 分散: $\frac{1}{n_j - 1} \sum_{i=1}^{n_j} \left(y_i^{(j)} - \bar{y}^{(j)} \right)^2 = w^T \left\{ \frac{1}{n_j - 1} \sum_{i=1}^{n_j} (x_i^{(j)} - \bar{x}_j)(x_i^{(j)} - \bar{x}_j)^T \right\} w$

$$= w^T S_j w$$

(2) $\sum_{j=1}^{g} n_j \left(\bar{y}^{(j)} - \bar{y} \right)^2 = w^T \left\{ \sum_{j=1}^{g} n_j (\bar{x}_j - \bar{x})(\bar{x}_j - \bar{x})^T \right\} w = w^T B w$

$$B := \sum_{j=1}^{g} n_j (\bar{x}_j - \bar{x})(\bar{x}_j - \bar{x})^T$$

(3) それぞれの群内の散らばりの程度は

$$\sum_{j=1}^{g} \sum_{i=1}^{n_j} \left(y_i^{(j)} - \bar{y}^{(j)} \right)^2 = w^T \left\{ \sum_{j=1}^{g} (n_j - 1) S_j \right\} w = w^T W w$$

$$W := \sum_{j=1}^{g} (n_j - 1) S_j$$